Human Identification and Recognition of Emotional State from Visual Input
D. Metaxas

Supplementary Report

The research through this grant has resulted in the following publications related to Deception:

1. Publications:

Book Chapters:

a. "Dynamic Data Driven Coupling of Continuous and Discrete Methods for 3D Tracking," D. Metaxas and G. Tsechpenakis. Lecture Notes in Computer Science, Springer-Verlag, vol. 3515, p. 712, April 2005.

Journal Publications:

a. "Optical Computer Recognition of Facial Expressions Associated with Stress Induced by Performance Demands". D. Dinges, R. Rider, J. Dorrian, E. McGlinchey, N. Rogers, Z. Cizman, S. Goldentein, C. Vogler, S. Venkataraman and D. Metaxas. Journal of Aviation, Space and Environmental Medicine, 2005

b. "Incremental Model-based Estimation Using Geometric Constraints", C. Sminchisescu, D. Metaxas and S. Dickinson. IEEE PAMI, 2005.

c. "Deception Detection through Automatic Unobtrusive Analysis of Nonverbal Behavior," T. O. Meservy, M. L. Jensen, J. Kruse, D. P. Twitchell, G. Tsechpenakis, J. K. Burgoon, D. N. Metaxas, and J. F. Nunamaker Jr." IEEE Intelligent Systems, Special Issue on Artificial Intelligence for National and Homeland Security, pp. 36-43, September/October 2005.

d. **"Statistical Cue Integration in DAG Deformable Models", S. Goldenstein, C. Vogler, and D. Metaxas. IEEE Pattern Analysis and Machine Intelligence, August 2003.**

Conference Publications

a. S. Lu, G. Tsechpenakis, D. Metaxas, M. L. Jensen, and J. Kruse, "Blob Analysis of the Head and Hands: A Method for Deception Detection and Emotional State Identification," Hawaii International Conference on System Sciences, Big Island, Hawaii, January 2005.

b. G. Tsechpenakis, D. Metaxas, M. Adkins, J. Kruse, M. Jensen, T. Meservy, D. Twitchell, A. Deokar and J. Nunamaker, "HMM-based Deception Recognition from Visual Cues," *IEEE International Conference on Multimedia & Expo (ICME'05)*, Amsterdam, The Netherlands, July 2005.

c. Tsechpenakis, D. Metaxas and S. Lu, "Analysis and Recognition of Nonverbal Behavior Using 2D Visual Cues and Machine Learning," Project Panel: Machine Learning Technology for Recognition and

Analysis of Suspicious Behavior from Human Gestures and Movement, *2nd Conference of the International Society for Gesture Studies (ISGS), Interacting Bodies*, Lyon, France, June 2005.

d. 3D Distinguishing Mislabeled Data from Correctly Labeled Data in Classifier Design. S. Venkataraman, D. Metaxas, D. Fradkin, C. Kulikowski, I. Muchnik. Procs ICTA, Nov 15-17, Boca Raton, Florida, 2004.
e. A Hybrid Face Recognition Method using Markov Random Fields. R.Huang, V. Pavlovic, and D. Metaxas.in Proceedings of ICPR, 2004.
f. "High Resolution Acquisition, Learning and Transfer of Dynamic 3-D Facial Expressions," Yang Wang, Xiaolei Huang, Chan-Su Lee, Song Zhang, Zhiguo Li, Dimitris Samaras, Dimitris Metaxas, Ahmed Elgammal, and Peisen Huang," In Proc. of the Annual Conf. of the European Association for Computer Graphics, Eurographics'04, Grenoble, France, August, 2004.
g. 3D Facial Tracking from Corrupted Movie Sequences. S. Goldenstein, C. Vogler and D. Metaxas. Procs. CVPR 2004.
h. Human Gait Recognition. R. Zhang, C. Vogler and D. Metaxas. Procs Articulated and Nonrigid Motion Conference (in conjunction with CVPR 2004). June 2004.
i. A Hierarchical Framework for High Resolution Facial Expression Tracking. X. Huang, S. Zhang, Y. Wang, D. Metaxas and D. Samaras. Procs Articulated and Nonrigid Motion Conference (in conjunction with CVPR 2004). June 2004.
j. Outlier Rejection in Deformable Model Tracking. S. Goldenstein, C. Vogler, J. Stolfi, V. Pavlovic and D. Metaxas. Procs Articulated and Nonrigid Motion Conference (in conjunction with CVPR 2004). June 2004.

2. Interactions/Transitions:

a. We have presented our work in all major conferences and related meetings (CVPR, ICCV), as well as the annual progress report meeting by Dr Herklotz. In addition I have given several distinguished lectures on this work at major Universities such as: Columbia University, Harvard and MIT.
b. Consultative and advisory functions to other laboratories and agencies: I am currently actively working with UPENN and Univ of Arizona on security issues.
c. Transitions. Our aim is to continue the development of this work and deploy it to airports, buildings, airplanes, embassies as well as in prisons. We are in the process of making contact to transition some of this technology for security purposes and possible deployment in Iraq and other security installations.

3. New discoveries, inventions, or patent disclosures: For the first time we have shown that it is possible to detect computationally from visual input normal from

abnormal gait as well as be able to focus on a person's gestures and facial expressions. We intend to patent some of our findings in the near future.

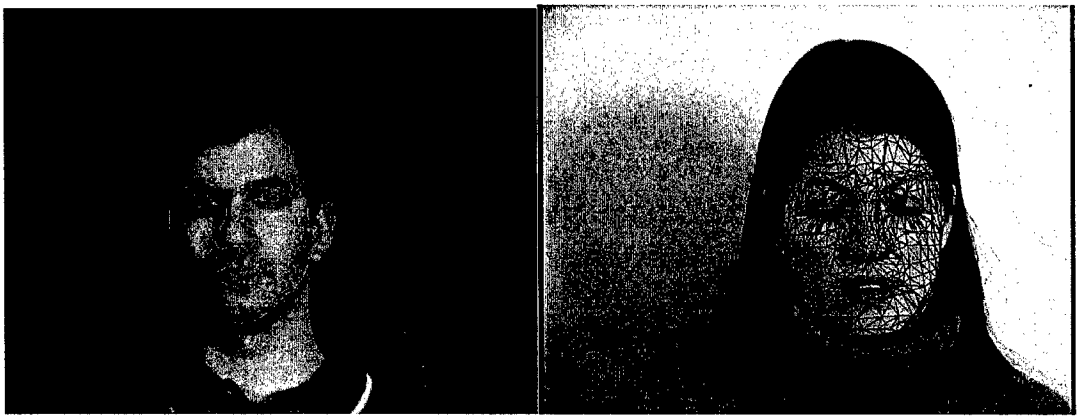
a. *Human body tracking and activity recognition*: (i) At a medium scale with cameras that can see details of people, we observe human kinesic behaviors (postures, head and trunk movements, and gestures) and proxemic behaviors (distancing and approach-avoidance). Since we are able to extract detailed body information in four dimensions (space and time) of people existing in a specific room and in a close range from the camera, we will be able to detect suspicious behaviors. To achieve this goal, we will need predefined criteria and features given from experts in the field of behavioral analysis. (ii) At a larger scale with broad-scope cameras, we can observe the whole-body motions of people within the surveillance area and track them using difference images with background and foreground modeling. Tracking multiple people enables us to detect unusual paths, velocities, and accelerations as clues to deception and in this way we will be able to build a model of human path selection, based on the movement style and content. Figure 1 shows two frames of a specific person being tracked in a crowd for human activity recognition.

b. *Face tracking and emotional state recognition based on facial features*: We have developed a fast and robust face detection technique for extracting faces from moving and/or cluttered backgrounds and under different lighting conditions. This is the first step, crucial for facial analysis. After a face region is detected in an image (the first frame of a short sequence), we have implemented a fusion between Kalman filter-based methods and ASM (Active Shape Models) to extract in 2D the facial features, while the individual is speaking or making facial expressions. Along with the 2D facial features extraction, the estimated facial points are used as input to our model-based 3D face tracker. The key contributions of the extracted 3D information are the automatic estimation of the exact head pose and the relative depth of the estimated features. In this way, we overcome the head pose limitations commonly met in existing applications. The 3D face tracking method provides robustness under strong head rotations/movements and occlusions. Moreover, the utilized deformable face model gives us information about 3D distances between specific facial points (features) and their motion (deformation) over time. Being able to extract the 3D facial information close to real-time, we can model our results in any format appropriate for facial expression recognition, emotional state recognition from facial features, and/or person identification from facial features. Figure 2 illustrates the facial features detection and tracking in 2D as well as the 3D face tracking result for an input frame.

c. *Head and hands blob-based tracking and emotional state recognition*: The first important step towards gesture analysis and emotional state recognition/deception detection from visual input is to detect and track the body parts of our interest, i.e. the head and hands. Although research efforts have investigated this issue in the recognition level, accurate and

real-time tracking of people and their body parts is still an open topic in the computer science community. Using color analysis, eigenspace-based shape segmentation, and Kalman filters, we have been able to track the position, size, and angle of different body parts with great accuracy and high rates close to 50 fps. Blob analysis extracts hand and face regions using the color distribution from an image sequence using prior knowledge from a skin color database. From the blobs, the left hand, right hand and face are tracked continuously, i.e. using the motion information over time. Also, from positions and movements of the hands and face we can make further inferences about the torso and the relation of each body part to other people and objects. Tracking the hands and the head, we are able to extract movement signatures, positions, velocities, accelerations and relative positions between the extracted blobs. We are able to automatically detect events such as when two hands come together and when a hand touches the face, and estimate these events' frequency and duration. For the recognition of two basic emotional states, i.e. "over-controlled" and "relaxed", we implemented a hierarchical recognition scheme, based on Hidden Markov Models, that uses as input visual cues and extracts a conclusion about the individuals behavior. For the construction of this recognition scheme, we used training and testing data including actors and real interrogation scenarios with ground truth. Figure 3 shows the blob extraction for the head and hands as well as shoulder detection results in an input frame.

Figure 1: Multiple people tracking: selecting and tracking a specific individual in a crowd



(a) (b)

Figure 2: (a) facial features extraction after the face is detected, (b) 3D face model fitting from 2D feature extraction.

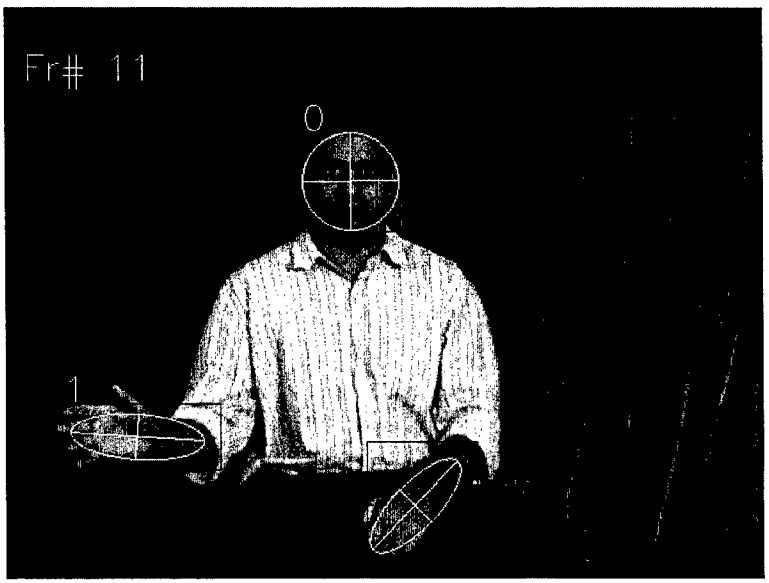


Figure 3: Head, hands and shoulders extraction

5) Personnel Supported

As part of this grant we were able to support:

a) **Gabriel Tsechpenakis Post Doctoral Fellow**
b) **Xiaolei Huan PhD student**
c) **Sundara Venkataraman PhD student**
d) **Summer support for D. Metaxas**

REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-06-0101

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
		12/15/02-12/14/05

4. TITLE AND SUBTITLE
Human Identification and Recognition of Emotional State from Visual Input

5a. CONTRACT NUMBER
F49620-03-1-0047

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
D. Metaxas

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Rutgers, The State Univ of NJ
Dept of Computer Science
110 Frelinghuysen Road
Piscataway, NJ 08854

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Rutgers, The State Univ of NJ
Office of Research &
Sponsored Programs
ASB III, 3 Rutgers Plaza
New Brunswick, NJ 08901 Dr Herklotz

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release, distribution unlimited

Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT We have developed a robust framework for analyzing human motion, for the purposes of identifying people from their movements, and recognizing the emotional state of people from their movements. The key point of this framework is the integration of data from multiple visual sources, such as gait, facial expressions, and body movements. In addition, the algorithms for integrating the data should be general enough to allow contributions from nonvisual sources, such as speech. This framework has immediate applications in the areas of surveillance and interrogation. In surveillance, we can detect intruders through people identification. In interrogation, we provide an invaluable backup for human interrogators and psychologists by picking up subtle behavioral cues from human motion that an interrogator might miss. By recognizing these cues, we can offer valuable cues to interrogators. Our system opens the way for the quantitative analysis of human communication in general. Collaborations: University of Arizona and UPENN on many of the above applications.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*